

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2004-205738

(43)Date of publication of application : 22.07.2004

(51)Int.Cl.

G10H 1/00

(21)Application number : 2002-373768

(71)Applicant : NAKAMURA SHUNSUKE

(22)Date of filing : 25.12.2002

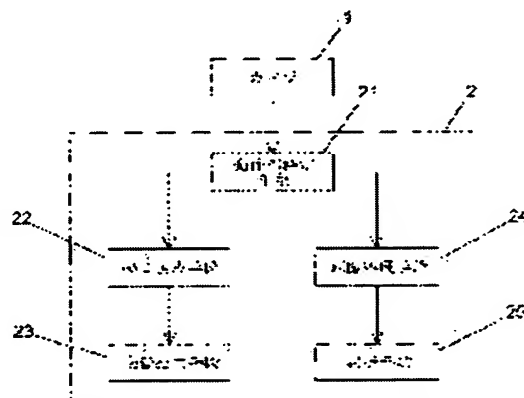
(72)Inventor : NAKAMURA SHUNSUKE

## (54) APPARATUS, PROGRAM, AND METHOD FOR MUSICAL SOUND GENERATION

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a musical sound generating apparatus that automatically generates musical sound data from inputted image data without preparing musical performance information etc.

**SOLUTION:** The apparatus is equipped with a camera 1 which continuously images a subject and inputs image data, frame by frame, an operating section specifying means 21 of specifying respective positions where the subject moves in frames from the image data inputted by the camera 1, a musical sound generating means 22 of generating musical sound data corresponding to the positions specified by the operating section specifying means 21, and a musical sound output means 23 of outputting the musical sound data generated by the musical sound generating means 22. The subject moves in front of the camera 1 and then a musical sound corresponding thereto is automatically generated and played with timbres of various musical instruments.



## LEGAL STATUS

[Date of request for examination] 25.12.2002

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3643829

[Date of registration] 04.02.2005

[Number of appeal against examiner's decision  
of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-205738

(P2004-205738A)

(43) 公開日 平成16年7月22日(2004.7.22)

(51) Int.Cl.<sup>7</sup>  
G10H 1/00F1  
G10H 1/00 102Zテーマコード(参考)  
5D378

審査請求 有 請求項の数 11 O L (全 9 頁)

(21) 出願番号 特願2002-373768 (P2002-373768)  
(22) 出願日 平成14年12月25日 (2002.12.25)(71) 出願人 503003304  
中村 俊介  
愛知県名古屋市南区星宮町293  
(74) 代理人 100099508  
弁理士 加藤 久  
(72) 発明者 中村 俊介  
愛知県名古屋市南区星宮町293  
Fターム(参考) 5D378 MM12 MM35 MM41

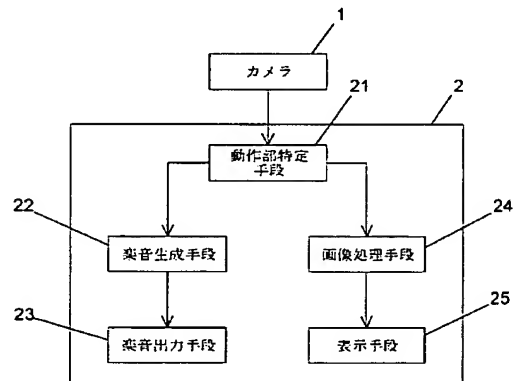
(54) 【発明の名称】 楽音生成装置、楽音生成プログラムおよび楽音生成方法

## (57) 【要約】

【課題】演奏情報などを予め準備することなく、入力した画像データから楽音データを自動生成する楽音生成装置の提供。

【解決手段】被写体を連続的に撮像し、フレームごとに画像データとして入力するカメラ1と、カメラ1により入力された画像データから、フレーム内で被写体の動きが有ったそれぞれの位置により特定する動作部特定手段21と、動作部特定手段21により特定された位置に応じた楽音データを生成する楽音生成手段22と、楽音生成手段22により生成された楽音データを出力する楽音出力手段23とを備え、カメラ1の前で被写体が動作することによって、その被写体の動作に応じた楽音が自動的に生成され、様々な楽器の音色で演奏される。

【選択図】 図2



**【特許請求の範囲】****【請求項 1】**

被写体を連続的に撮像し、フレームごとに画像データとして入力する撮像手段と、  
同撮像手段により入力された画像データから、フレーム内で前記被写体の動きが有ったそれぞれの位置により特定する動作部特定手段と、  
同動作部特定手段により特定された位置に応じた楽音データを生成する楽音生成手段と、  
同楽音生成手段により生成された楽音データを出力する手段と  
を備えた楽音生成装置。

**【請求項 2】**

前記動作部特定手段は、前記位置の特定を、複数のフレーム間の画像データの比較により 10  
行うものである請求項 1 記載の楽音生成装置。

**【請求項 3】**

前記楽音生成手段は、前記動作部特定手段により特定された位置に応じた楽器の音源により前記楽音データを生成するものである請求項 1 または 2 に記載の楽音生成装置。

**【請求項 4】**

前記楽音生成手段は、前記動作部特定手段により特定された位置に応じた音階により前記楽音データを生成するものである請求項 1 から 3 のいずれかに記載の楽音生成装置。

**【請求項 5】**

前記楽音生成手段は、前記動作部特定手段により特定された位置に応じた音量バランスにより前記楽音データを生成するものである請求項 1 から 4 のいずれかに記載の楽音生成装置。 20

**【請求項 6】**

前記撮像手段により入力された画像データを、前記動作部特定手段により特定された位置に応じて画像処理する画像処理手段と、  
同画像処理手段により画像処理された画像データを表示する表示手段と  
を備えた請求項 1 から 5 のいずれかに記載の楽音生成装置。

**【請求項 7】**

前記画像処理手段は、前記動作部特定手段により特定された位置に応じて前記画像データの配色を変えるものである請求項 6 記載の楽音生成装置。

**【請求項 8】**

被写体を連続的に撮像し、フレームごとに画像データとして入力する撮像手段が接続されたコンピュータを、  
前記撮像手段により入力された被写体の画像データから、フレーム内で前記被写体の動きが有ったそれぞれの位置により特定する動作部特定手段と、  
同動作部特定手段により特定された位置に応じた楽音データを生成する楽音生成手段と、  
同楽音生成手段により生成した楽音データを出力する手段  
として機能させるための楽音生成プログラム。 30

**【請求項 9】**

さらに前記コンピュータを、  
前記撮像手段により入力された画像データを、前記動作部特定手段により特定された位置 40  
に応じて画像処理する画像処理手段  
として機能させるための請求項 8 記載の楽音生成プログラム。

**【請求項 10】**

被写体を連続的に撮像し、フレームごとに画像データとして入力し、この入力された画像データから、フレーム内で前記被写体の動きが有ったそれぞれの位置により特定し、この特定された位置に応じた楽音データを生成し、この生成された楽音データを出力する楽音生成方法。

**【請求項 11】**

さらに、前記フレームごとに入力された画像データを、前記特定された位置に応じて画像処理し、この処理された画像データを表示する請求項 10 記載の楽音生成方法。 50

**【発明の詳細な説明】****【0001】****【発明の属する技術分野】**

本発明は、画像データに対応した楽音データを自動生成する楽音生成装置、楽音生成プログラムおよび楽音生成方法に関する。

**【0002】****【従来の技術】**

画像に対応して演奏の制御を行う技術として、例えば特許文献1には、被写体の輪郭を利用してテンポ等を制御する技術が開示されている。この技術では、入力されたビデオ信号からR（赤）、G（緑）、B（青）の各色信号を分離し、各色毎にデジタルデータとして階調を表す階調データを生成する。そして、各色の階調データと予め定めた閾値データとに基づいて被写体を特定し、当該被写体の輪郭を検出し、この検出した輪郭の複雑さに応じて演奏を制御する。

10

**【0003】**

しかしながら、この特許文献1に開示の技術では、被写体を特定し、輪郭を検出する必要があるため、処理にかかる負担が大きいという問題がある。そこで、この問題を解決するための技術として、例えば特許文献2には、供給された画像から複数の動きベクトルを抽出し、この抽出した複数の動きベクトルから一つの制御ベクトルを算出し、この算出した制御ベクトルに基づいて、演奏を制御する技術が開示されている。

**【0004】**

20

**【特許文献1】**

特許第2629740号公報

**【特許文献2】**

特開2000-276139号公報

**【0005】****【発明が解決しようとする課題】**

上記特許文献1、2記載の技術は、いずれも別途供給された演奏の内容を示す演奏情報等から楽音を再生する際に、ビデオ画像に応じた制御を行うことにより、楽音にアレンジを加えるものである。すなわち、これらの技術は演奏情報が何もない状態から楽音を生成するものではない。したがって、これらの技術を利用するためには、予め演奏情報とビデオ画像の両方をそれぞれ準備しておく必要がある。

30

**【0006】**

そこで、本発明においては、演奏情報などを予め準備することなく、入力した画像データから楽音データを自動生成する楽音生成装置、楽音生成プログラムおよび楽音生成方法を提供する。

**【0007】****【課題を解決するための手段】**

本発明の楽音生成装置は、被写体を連続的に撮像し、フレームごとに画像データとして入力する撮像手段と、撮像手段により入力された画像データから、フレーム内で被写体の動きが有ったそれぞれの位置により特定する動作部特定手段と、動作部特定手段により特定された位置に応じた楽音データを生成する楽音生成手段と、楽音生成手段により生成された楽音データを出力する手段とを備えたものである。

40

**【0008】**

本発明の楽音生成装置によれば、撮像手段により連続的に撮像され、フレームごとに入力された画像データに基づいて、動作部特定手段により、フレーム内で被写体の動きが有ったそれぞれの位置が特定される。そして、楽音生成手段により、この特定されたフレーム内の位置に応じた楽音データ、すなわち被写体の動く位置に応じた楽音データが生成され、出力手段により出力される。

**【0009】**

動作部特定手段は、位置の特定を、複数のフレーム間の画像データの比較により行うもの

50

とするのが望ましい。複数のフレーム間の画像データを比較することにより、複数のフレーム間の画像データに変化があれば、その複数のフレーム間で被写体に動きが有ったことになるため、その変化があった位置を被写体に動きが有った部分として容易に特定することができる。

【0010】

楽音生成手段は、動作部特定手段により特定された位置に応じた楽器の音源により楽音データを生成するものであることが望ましい。これにより、被写体の動く位置に応じて、それぞれ異なる楽器の音源により生成した楽音データを得ることができる。

【0011】

楽音生成手段は、動作部特定手段により特定された位置に応じた音階により楽音データを生成するものであることが望ましい。これにより、被写体の動く位置に応じて、それぞれ異なる音階により生成した楽音データを得ることができる。

10

【0012】

楽音生成手段は、動作部特定手段により特定された位置に応じた音量バランスにより楽音データを生成するものであることが望ましい。これにより、被写体の動く位置に応じて、音量バランスを調整した楽音データを得ることができる。

【0013】

また、本発明の楽音生成装置は、撮像手段により入力された画像データを、動作部特定手段により特定された動作部特定手段により特定された位置に応じて画像処理する画像処理手段と、画像処理手段により画像処理された画像データを表示する表示手段とを備えたものとするのが望ましい。例えば、画像処理手段は、動作部特定手段により特定された位置に応じて画像データの配色を変えるものとすることができる。

20

【0014】

これにより、被写体の動く位置に応じて生成される楽音データに加えて、被写体の映像が、そのまま表示手段に表示されるのではなく、その被写体の動く位置に応じて配色が変えられる等のアレンジが施されて表示される。

【0015】

上記本発明の楽音生成装置は、被写体を連続的に撮像し、フレームごとに画像データとして入力する撮像手段が接続されたコンピュータを、撮像手段により入力された被写体の画像データから、フレーム内で前記被写体の動きが有ったそれぞれの位置により特定する動作部特定手段と、動作部特定手段により特定された位置に応じた楽音データを生成する楽音生成手段と、楽音生成手段により生成した楽音データを出力する手段として機能させるための楽音生成プログラムの実行により実現可能である。

30

【0016】

また、この楽音生成プログラムは、さらにコンピュータを、撮像手段により入力された画像データを、動作部特定手段により特定された位置に応じて画像処理する画像処理手段として機能させるためのものとすることもできる。

【0017】

【発明の実施の形態】

図1は本発明の実施の形態における楽音生成装置の斜視図、図2は図1の楽音生成装置の構成を示すブロック図である。

40

【0018】

図1および図2に示すように、本発明の実施の形態における楽音生成装置は、被写体を撮像するための撮像手段としてのカメラ1と、このカメラ1が接続され、カメラ1により入力された画像データに基づいて各種処理を行うコンピュータ2とにより構成される。カメラ1は、被写体を例えば1秒間に8フレーム連続的に撮像し、フレームごとに画像データとしてコンピュータ2へ入力するものである。

【0019】

コンピュータ2は、カメラ1により入力された画像データからそれぞれのフレーム内で被写体の動きが有った部分を特定する動作部特定手段21と、動作部特定手段21により特

50

定されたフレーム内の部分に応じた楽音データを生成する楽音生成手段 2 2 と、楽音生成手段 2 2 により生成された楽音データを出力する楽音出力手段 2 3 と、カメラ 1 により入力された画像データを画像処理する画像処理手段 2 4 と、画像処理手段 2 4 により画像処理された画像データを表示する表示手段 2 5 として機能する。

#### 【 0 0 2 0 】

動作部特定手段 2 1 は、カメラ 1 により入力された画像データから、フレーム内で被写体の動きが有ったそれぞれの位置を特定するものである。例えば、動作部特定手段 2 1 は、この被写体の動きがあった位置を予め分割されたフレーム内の区画の位置により特定する。図 3 および図 4 はフレームの分割例を示している。図 3 は横方向の分割例であり、横方向の区画を A、B、C、D、E の 5 分割としている。図 4 は区画 A の縦方向の分割例であり、縦方向の区画 A に割り当てる楽器の音域を音階ごとに分割したものである。区画 B ～ E の縦方向についても同様に各区画 B ～ E に割り当てる楽器の音域を音階ごとに分割する。

10

#### 【 0 0 2 1 】

動作部特定手段 2 1 は、このフレーム内の縦横の区画の位置の特定を、複数のフレーム間の画像データの比較により行う。図 5 は、動作部特定手段 2 1 による動きの判定の様子を示している。図 5 に示すように、動作部特定手段 2 1 は、カメラ 1 により入力された画像データから、現在のフレームの画像とその直前のフレームの画像（または背景画像）のすべての画素（ピクセル）の明るさを比較し、その差が既定値以上ならば動き有りの画素であると判断する。このとき、動作部特定手段 2 1 は、動き有りの画素群の重心（X 座標、Y 座標）が縦横のどの区画に属するものであるかを判定する。

20

#### 【 0 0 2 2 】

楽音生成手段 2 2 は、動作部特定手段 2 1 により特定されたフレーム内の縦横の区画の位置に応じた楽音データ（例えば、MIDI（Musical Instrument Digital Interface）規格のデータ）を生成するものである。図 3 に示すように、フレームが横方向に 5 分割されている場合、楽音生成手段 2 2 は、これらの区画 A ～ E の中から被写体の動きが有った部分に対応する楽器（例えば、A：ピアノ、B：ギター、C：ベース、D：ドラム、E：オルゴール等。）の音源により楽音データを生成する。

#### 【 0 0 2 3 】

また、楽音生成手段 2 2 は、各区画 A ～ E 中の縦方向の区画の中から被写体の動きがあった部分に対応する音階の音により楽音データを生成する。このとき、楽音生成手段 2 2 は、基準となるコードを設定し、音楽理論に基づいてそのコード上の音階のみを選択する。例えば、コードが「C」の場合は、図 6 に示すように基準コード（ベース）は 1 2 の倍数で、これに 0、4、7 を加えた値を音階（ノートナンバー）とする。なお、発音のタイミングは、毎回発音、2 回ごとに発音、2 回発音して 4 回休むなど、各楽器に応じて設定する。また、打楽器については特殊であるため、ノートナンバーは音階ではなく鳴らす打楽器の種類とする。すなわち、縦方向の動きの変化は打楽器の種類の変化となる。

30

#### 【 0 0 2 4 】

なお、基準コードは時間の経過や何らかの動作をきっかけ（トリガー）として変化させる。この基準コードの変化（コード進行）によって音楽として美しく奏でられるようになる。この変化には、予め順番やタイミング等を指定しておく方法や、ランダムで進行させる方法などがある。変化のタイミングとしては、時間の経過による変化の他、被写体の動きの量によって切り替えるなど、被写体の動きに合わせて変化させることも可能である。

40

#### 【 0 0 2 5 】

さらに、楽音生成手段 2 2 は、フレーム内の区画の位置に応じた音量バランスにより楽音データを生成する。図 7 は音量バランスの設定例を示している。楽音生成手段 2 2 は、図 7 に示すように、動作部特定手段 2 1 により特定されたフレーム内の横の位置に応じた音量バランスで、後述する左右のスピーカ 3 a、3 b から音出力されるように楽音データを生成する。

#### 【 0 0 2 6 】

50

図 2 に戻って、楽音出力手段 2 3 は、本実施形態においては、図 1 に示すようにコンピュータ 2 の左右に配置され、それぞれ楽音生成手段 2 2 により生成された楽音データを音として出力するスピーカ 3 a, 3 b とする。なお、楽音出力手段 2 3 はスピーカ 3 a, 3 b のように直接音を出力するものでなくてもよい。例えば、楽音出力手段 2 3 は、外部接続される音源（外部 M I D I 音源等）により音を出力するための音データを出力するものであってもよい。

#### 【 0 0 2 7 】

画像処理手段 2 4 は、カメラ 1 により入力された画像データを、動作部特定手段 2 1 により特定された被写体の動き位置に応じて配色を変えることにより画像処理するものである。例えば、被写体の動きが有った部分だけ着色したり、動きが有った部分を着色して上から重ねて塗ったり、着色する色や順番を変えたりする。

#### 【 0 0 2 8 】

表示手段 2 5 は、本実施形態においては、コンピュータ 2 の正面に配置された内蔵ディスプレイ 4 とするが、外部接続されるディスプレイ、C R T（陰極線管）、プロジェクタ、テレビジョン等であってもよい。

#### 【 0 0 2 9 】

以下、上記構成の楽音生成装置による楽音生成手順について、図 8 のフロー図に基づき説明する。

#### 【 0 0 3 0 】

（S 1 0 1）初期設定として、生成する楽音のテンポ、楽器の数と種類、音楽コード進行、色の変化などの設定を行う。

（S 1 0 2）カメラ 1 からコンピュータ 2 へ現在の画像データを入力する。

（S 1 0 3）動作部特定手段 2 1 により、動きを抽出（被写体の動きが有ったそれぞれの位置を特定）する。

#### 【 0 0 3 1 】

（S 1 0 4）音処理では、楽音生成手段 2 2 により被写体の動きの位置と変化量を算出する。

（S 1 0 5）被写体の動きの変化量が大きい場合には、コード進行に変化を付ける。なお、変化量が極小で続く場合には、動きがなくなつたと判断（S 1 0 8）し、（S 1 0 1）へ戻る。

（S 1 0 6）楽音生成手段 2 2 は、被写体の動きの位置と変化量から楽音データ（M I D I 規格データ）を生成する。このとき、楽音生成手段 2 2 は、被写体の動きの変化量に応じて楽音データの音量を決定する。

（S 1 0 7）生成された M I D I 規格データは、楽音出力手段 2 3 へ送り、スピーカ 3 a, 3 b から出力する。

#### 【 0 0 3 2 】

（S 1 0 9）画像処理では、カメラ 1 から入力された画像データの、画像処理手段 2 4 により被写体の動きが有った部分の画素（ピクセル）を指定された色で塗る。

（S 1 1 0）画像処理手段 2 4 により画像処理された画像データをディスプレイ 4（表示手段 2 5）へ表示する。

（S 1 0 2）から（S 1 1 0）のステップは、例えば 1 秒間に 8 回、初期設定により設定したテンポの速さで繰り返す。なお、このテンポは、上記処理の途中で被写体の動きに応じて変化させる構成とすることも可能である。

#### 【 0 0 3 3 】

以上のように、本実施形態における楽音生成装置によれば、コンピュータ 2 に接続されたカメラ 1 により入力した画像データから楽音データが自動生成され、スピーカ 3 a, 3 b などにより出力される。すなわち、従来のように演奏情報などを予め準備することなく、コンピュータ 2 に接続されたカメラ 1 の前で被写体が動作をすることによって、その被写体の動作に応じた楽音が自動的に生成され、様々な楽器の音色で演奏される。

#### 【 0 0 3 4 】

10

20

30

40

50



また、本実施形態における楽音生成装置によれば、楽音の演奏とともに、カメラ 1 により入力された画像データが、被写体の動きに応じてアレンジされて、コンピュータ 2 のディスプレイ 4 に表示される。したがって、本実施形態における楽音生成装置は、生成された楽音を聴覚だけでなく視覚でも楽しむことができ、音と映像を合わせて様々な分野で利用することが可能である。

#### 【 0 0 3 5 】

このような本実施形態における楽音生成装置は、例えば表 1 に挙げるような場面で利用することが可能である。

【表 1】

利用方法	対象者
楽器・パフォーマンスツール	音楽家・楽器に興味のある人
教育ツール（音・色）	子供
リハビリテーション用ツール	怪我や病気の人
医療用ツール（精神科・老人）	自閉症などの精神病の人・老人
エンターテインメント（玩具・イベント）	子供・若者
コミュニケーションツール	視覚障害者等
店舗等のディスプレイ	ディスプレイ・イベント会社等

10

20

#### 【 0 0 3 6 】

なお、本実施形態においては、楽音生成装置をカメラ 1 と汎用のコンピュータ 2 により構成する例について説明したが、本発明の楽音生成装置はコンピュータ 2 を楽音生成の専用機やゲーム機等により置き換えて構成することも可能である。カメラ 1 については、外部接続の他、機器に内蔵した構成とすることが可能である。また、カメラ 1 を複数台接続して複数の画像データを同時に入力し、同時演奏させることも可能である。このとき、カメラ 1 を複数の離れた場所にそれぞれ設置し、ネットワークを通じてコンピュータ 2 に接続することにより、複数の場所からの同時演奏を行わせることも可能である。

30

#### 【 0 0 3 7 】

また、動作部特定手段 2 1 によるフレームの分割については、本実施形態における分割方向を縦横入れ替えて構成することも可能である。さらに、縦横を組み合わせたリ、斜め方向、半径方向、円周方向に分割したりすることも可能である。

#### 【 0 0 3 8 】

##### 【発明の効果】

本発明により、以下の効果を奏することができる。

#### 【 0 0 3 9 】

(1) 被写体を連続的に撮像し、フレームごとに画像データとして入力し、この入力された画像データから、フレーム内で被写体の動きが有ったそれぞれの位置により特定し、この特定された位置に応じた楽音データを生成し、この生成された楽音データを出力する構成により、従来のように演奏情報などを予め準備することなく、被写体が動作をすることによって、その被写体の動作に応じた楽音が自動的に生成され、出力される。これにより、被写体の動作だけから自動的に生成された楽音を楽しむことが可能となる。

40

#### 【 0 0 4 0 】

(2) さらに、フレームごとに入力された画像データを、特定された位置に応じて画像処理し、この処理された画像データを表示する構成により、被写体の動作に応じてアレンジされた被写体の画像が出力される。これにより、生成された楽音を映像でも楽しむことが

50

可能となる。

【図面の簡単な説明】

【図 1】本発明の実施の形態における楽音生成装置の斜視図である。

【図 2】図 1 の楽音生成装置の構成を示すブロック図である。

【図 3】フレームの横方向の分割例を示す図である。

【図 4】フレームの縦方向の分割例を示す図である。

【図 5】動き判定の様子を示す図である。

【図 6】音楽理論に基づいた音階設定の例を示す図である。

【図 7】音量バランスの設定例を示す図である。

【図 8】本実施形態における楽音生成手順を示すフロー図である。

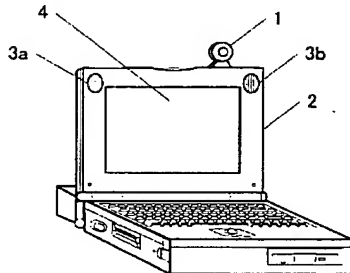
【符号の説明】

- 1 カメラ
- 2 コンピュータ
- 3 a, 3 b スピーカ
- 4 ディスプレイ
- 2 1 動作部特定手段
- 2 2 楽音生成手段
- 2 3 楽音出力手段
- 2 4 画像処理手段
- 2 5 表示手段

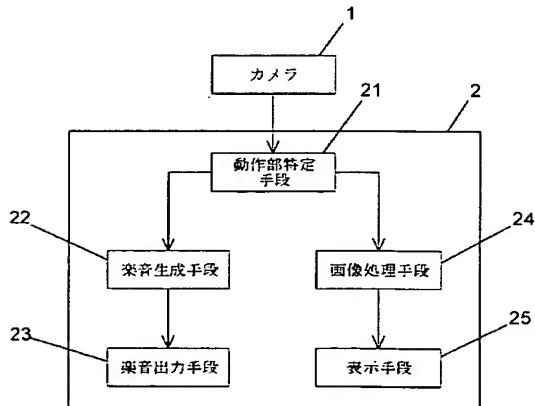
10

20

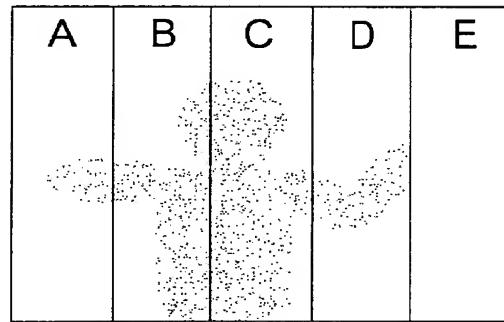
【図 1】



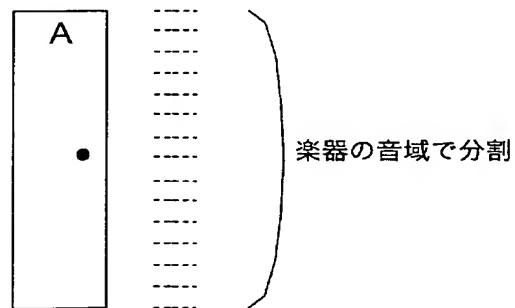
【図 2】



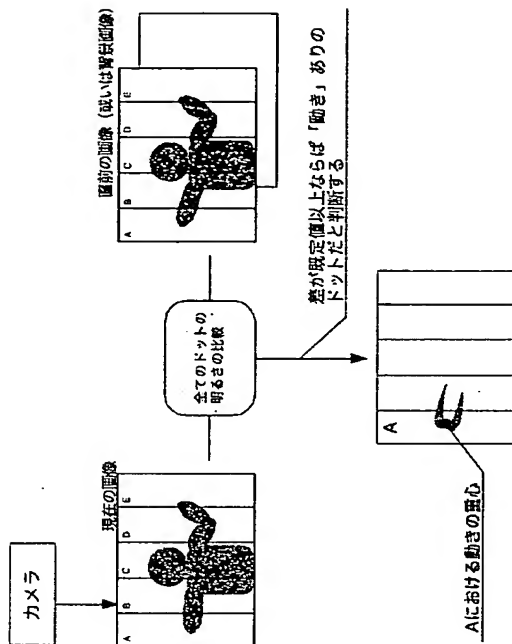
【図 3】



【図 4】



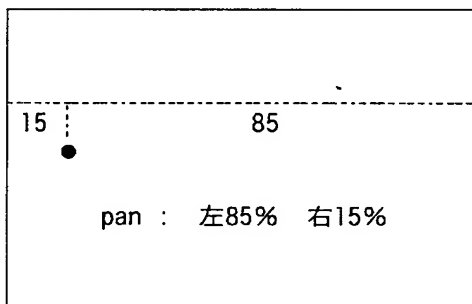
【図 5】



【図 6】

----	84	ベース (+0)
----	79	(+7)
----	76	(+4)
----	72	ベース (-0)
●----	67	(+7)
----	64	(+4)
----	60	ベース (-0)
----	55	(+7)
----	52	(+4)
----	48	ベース (+0)
----	44	(+7)
----	40	(+4)
----	36	ベース (-0)
----	32	(+7)
----	28	(+4)
----	24	ベース (-0)
----	20	(+7)
----	16	(+4)
----	12	ベース (-0)
----	8	(+7)
----	4	(+4)
----	0	ベース (+0)

【図 7】



【図 8】

